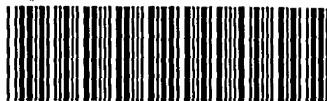


U30
RECEIVED

16A5



JUL 23 2002

TECH CENTER 1600/2900

OIPE

410
8-7-02
P2

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/045,594B

DATE: 07/02/2002
TIME: 10:16:38

Input Set : A:\DC-0134\DC-0134.app
Output Set: N:\CRF3\07022002\J045594B.raw

ENTERED

3 <110> APPLICANT: Taylor, Ronald K.
4 LaPointe, Christian F.
5 Trustees of Dartmouth College
7 <120> TITLE OF INVENTION: COMPOUNDS AND METHODS OF IDENTIFYING COMPOUNDS WHICH
8 INHIBIT A NEW CLASS OF ASPARTYL PROTEASES
10 <130> FILE REFERENCE: DC-0134
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/045,594B
C--> 13 <141> CURRENT FILING DATE: 2002-06-17
15 <150> PRIOR APPLICATION NUMBER: 60/143,355
16 <151> PRIOR FILING DATE: 1999-07-12
18 <160> NUMBER OF SEQ ID NOS: 4
20 <170> SOFTWARE: PatentIn Ver. 2.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 705
24 <212> TYPE: DNA
25 <213> ORGANISM: Staphylococcus aureus
27 <400> SEQUENCE: 1
28 ttggtagtat tgttatccta tagttcagt tgtattttta gttttctata tcaatttata 60
29 tctatagaag aaacgtcatt tgattattta catagaagat cgaaatgtga ttattttaat 120
30 tcgtcactca aatgttatga attaatgcgg attatttagtt ttttattttttaaaaaggcgaa 180
31 tgcgaaact gtcgaaagcg tatttcccta acacatttct taggggaaac ctttgctta 240
32 atacatatcg tctttattaa gtatgatttc acatacgtaa atgctacgtt atttataact 300
33 acatatgttt ttctgtttat ttttactatg accgatatac cttcttttaat gtttagattgt 360
34 cgcttaatata taattttatttgc tatcgtttct ctctcgtaa gtatgattta tccagtagct 420
35 ttatcatta ttatgtatgac cacgcatac ttctactttt tatttcgggc atatattgg 480
36 tatggtgacg ttttactaat atctgcactt tctttgtttt tccctctcoa attcactatt 540
37 tatgtcattt tatttacatt tgcattgtt ggttttagttt ctttaatttac catgatatttt 600
38 aagccgatata aactattacc ccttgttcca ttatattta tttccctttttt tatcaattca 660
39 cttttttata atgatatcca tcaattttta ggaggcgat atttt 705
42 <210> SEQ ID NO: 2
43 <211> LENGTH: 235
44 <212> TYPE: PRT
45 <213> ORGANISM: Staphylococcus aureus
47 <400> SEQUENCE: 2
48 Leu Val Val Leu Leu Ser Tyr Ser Cys Ser Cys Ile Phe Ser Phe Leu
49 1 5 10 15
51 Tyr Gln Phe Ile Ser Ile Glu Glu Thr Ser Phe Asp Tyr Leu His Arg
52 20 25 30
54 Arg Ser Lys Cys Asp Tyr Cys Asn Ser Ser Leu Lys Trp Tyr Glu Leu
55 35 40 45
57 Met Pro Ile Ile Ser Phe Leu Leu Leu Lys Gly Arg Cys Arg Asn Cys
58 50 55 60
60 Arg Lys Arg Ile Ser Leu Thr His Phe Leu Gly Glu Thr Phe Ala Leu

Input Set : A:\DC-0134\Dc-0134.app
Output Set: N:\CRF3\07022002\J045594B.raw

61	65	70	75	80
63	Ile Pro Ile Val Phe Ile Lys Tyr Asp Phe Thr Tyr Val Asn Ala Thr			
64	85	90	95	
66	Leu Phe Ile Thr Thr Tyr Val Phe Leu Leu Ile Phe Thr Met Thr Asp			
67	100	105	110	
69	Ile Thr Ser Leu Met Leu Asp Cys Arg Leu Ile Ile Tyr Cys Ile			
70	115	120	125	
72	Val Ser Leu Ser Leu Ser Met Ile Tyr Pro Val Ala Phe Ile Ile Ile			
73	130	135	140	
75	Ser Met Thr Thr His Ile Phe Tyr Phe Leu Phe Arg Ala Tyr Ile Gly			
76	145	150	155	160
78	Tyr Gly Asp Val Leu Leu Ile Ser Ala Leu Ser Leu Phe Phe Pro Leu			
79	165	170	175	
81	Gln Phe Thr Ile Tyr Val Ile Leu Phe Thr Phe Val Ile Ala Gly Leu			
82	180	185	190	
84	Val Ala Leu Ile Thr Met Ile Phe Lys Pro Ile Lys Leu Leu Pro Leu			
85	195	200	205	
87	Val Pro Phe Ile Phe Ile Ser Phe Phe Ile Asn Ser Leu Phe Tyr Asn			
88	210	215	220	
90	Asp Ile His Gln Phe Leu Gly Gly Val Tyr Phe			
91	225	230	235	
94	<210> SEQ ID NO: 3			
95	<211> LENGTH: 986			
96	<212> TYPE: DNA			
97	<213> ORGANISM: Vibrio cholerae			
99	<400> SEQUENCE: 3			
100	gaattcaaaa taaaacagat tgcataatcgag tctgggtttg ctgcgttcc atattttct 60			
101	acagtattta agtcaaccat gaatgttagca ccaagtgaat atttattttat gttgacagga 120			
102	gttgcagaaa aataatggaa tacgtttact tgatccattt ttgcattgtt tcactgattt 180			
103	tgggtatgtt tagaatgtt gtgatttatac gtctaccaag aaaaattttt ctaaagaatc 240			
104	atttttttta tgacatcgat tcaaattcgat ccattgtgcc taaatgtgga aataaaatca 300			
105	gttggatatga taatgttcca ttgttaagtt atttgttact tcatggtaag tgttagacatt 360			
106	gtgatgaaaa aatttcatta agttattttt ttgtcgagtt atcattttttt ataatcgat 420			
107	ttccaatata ttggttatca actgactggg ttgattcgat cgtattgctc gggctttact 480			
108	ttattttttta caatctttt gtaattgatt ttaaatctat gctattgcca aatttactta 540			
109	cgtatccgat attcatgttg gcatttattt atgtcgagcc aaatcaagca ttaacagttg 600			
110	aaagctctat tattgggtgtt ttgcggctt tcatcattac ctatgtgagt aatttcattt 660			
111	ttcggttatt caaaaagaatc gatgtgtatgg gaggtggaga tataaagctc tataactgca 720			
112	ttggtaactct tattcggtttt gagtttgatc catattttt cttgtgtcc tcgattattt 780			
113	catttattca ttggttttttt gctagatgtt catgtcggtt ttgcattat atccctcttt 840			
114	gaccaaggcat tattatctct ttgttaattt ttttttttca aatccgttta atgtgagcta 900			
115	aaatgaagtt ttctatcaag ttactaatga tttttacatc aattataatt actatttcat 960			
116	ctatccttac ttactttcag atgcat			986
119	<210> SEQ ID NO: 4			
120	<211> LENGTH: 253			
121	<212> TYPE: PRT			
122	<213> ORGANISM: Vibrio cholerae			
124	<400> SEQUENCE: 4			
125	Met Glu Tyr Val Tyr Leu Leu Phe Ser Ile Val Ser Leu Ile Leu			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/045,594B

DATE: 07/02/2002
TIME: 10:16:38

Input Set : A:\DC-0134\DC-0134.app
Output Set: N:\CRF3\07022002\J045594B.raw

126 1 5 10 15
128 Gly Ser Phe Ser Asn Val Val Ile Tyr Arg Leu Pro Arg Lys Ile Leu
129 20 25 30
131 Leu Lys Asn His Phe Phe Tyr Asp Ile Asp Ser Asn Arg Ser Met Cys
132 35 40 45
134 Pro Lys Cys Gly Asn Lys Ile Ser Trp Tyr Asp Asn Val Pro Leu Leu
135 50 55 60
137 Ser Tyr Leu Leu His Gly Lys Cys Arg His Cys Asp Glu Lys Ile
138 65 70 75 80
140 Ser Leu Ser Tyr Phe Ile Val Glu Leu Ser Phe Phe Ile Ile Ala Phe
141 85 90 95
143 Pro Ile Tyr Trp Leu Ser Thr Asp Trp Val Asp Ser Phe Val Leu Leu
144 100 105 110
146 Gly Leu Tyr Phe Ile Leu Phe Asn Leu Phe Val Ile Asp Phe Lys Ser
147 115 120 125
149 Met Leu Leu Pro Asn Leu Leu Thr Tyr Pro Ile Phe Met Leu Ala Phe
150 130 135 140
152 Ile Tyr Val Gln Pro Asn Gln Ala Leu Thr Val Glu Ser Ser Ile Ile
153 145 150 155 160
155 Gly Gly Phe Ala Ala Phe Ile Ile Thr Tyr Val Ser Asn Phe Ile Val
156 165 170 175
158 Arg Leu Phe Lys Arg Ile Asp Val Met Gly Gly Asp Ile Lys Leu
159 180 185 190
161 Tyr Thr Ala Ile Gly Thr Leu Ile Gly Val Glu Phe Val Pro Tyr Leu
162 195 200 205
164 Phe Leu Leu Ser Ser Ile Ile Ala Phe Ile His Trp Phe Phe Ala Arg
165 210 215 220
167 Val Ser Cys Arg Tyr Cys Leu Tyr Ile Pro Leu Gly Pro Ser Ile Ile
168 225 230 235 240
170 Ile Ser Phe Val Ile Val Phe Phe Ser Ile Arg Leu Met
171 245 250

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/045,594B

DATE: 07/02/2002
TIME: 10:16:39

Input Set : A:\DC-0134\DC-0134.app
Output Set: N:\CRF3\07022002\J045594B.raw

L:12 M:270 C: Current Application Number differs, Replaced Application Number
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date